智力科学的一些入门研究问题
 作者：顾立帆
一、引言
近年，人工智能在围棋、无人汽车、自然语言处理、制药等等研究方面有些进展。但缺乏准确、严谨的分析、测试方法，使这些进展被明显夸大，可能误导未来
的发展方向。
计算机围棋、无人汽车、自然语言处理跟人类智力密切交互。需要对高度不稳定的人类智力作结构性、系统性研究，才能有效设计科学实验、检测这些人工智能
技术、找出问题。
目前药物研发不得不较多使用猜测。在人类研究者所作前期研究已积累大量有效数据的特殊情况下，利用人类知识的人工智能可能比人类猜得更准。但在药物成
败的关键：漫长的临床试验上，人工智能未必有多少帮助。一旦人类在药物原理上取得重要进展，人工智能在中期的猜测作用也会消失。所以，需要新科学对人
类特有智力的潜力和未来发展作充分研究。
物理科学研究前沿已接近测量极限。近年生命科学，特别是智力研究，出现较多实验可重复性等等问题，部分原因跟观察量选择不当有关，也跟不正确地使用描
述语言和数学有关。计算机，包括量子计算机，缺乏足够抽象程度，无法解决逻辑和诡辩问题。诡辩引发决策错误更是危险。
生命科学只能研究演化，无法研究人类特有智力的进化和退化。由于熵定律的影响，生命实际上在逆水行舟。如果对人类特有智力的研究误入歧途，就会出现智
力的退化和文明的衰落。
很多人其实陷在牛顿世界观和拉普拉斯妖的错误观念中，Universal Approximation Theorem 仍延续这种错误。所以需要研究自由
意志是否存在，如果存在跟量子随机性、基因、自我意识有何本质区别，引入自由意志对人类直觉和智力模型的逻辑、结构、功能又有什么影响。
哥德尔指出的传统逻辑学问题，在人工智能和智力科学研究中会变得越来越严重。笔者分析了基于归纳和演绎的传统逻辑学缺陷，指出了逻辑排中律虽不具普适
性、但寻找满足逻辑排中律的逻辑临界点及其演变规律是智力进化的关键，以此建立更有效、更严谨的新逻辑体系。
Nash Embedding Theorems 有局限性，无法作为逻辑临界点分析的数学基础。所以计算机语言处理中的 Word-embedding Vector Space 无法处理关键的逻辑和诡辩问题，语言智力研究需要新的数学理论和方法。由于缺乏对语言智力的系统、结构研
究，现有测试计算机语言处理的数据集也无法测试关键的逻辑和诡辩问题和其它很多重要问题，大大高估了人工智能的智力水准。
笔者的智力科学理论可以证明 AlphaGo 系列缺乏人类直觉造成的问题、SAE level 4 自动驾驶不可能真正实现、Turing Test 无效，在
人工智能测试、教育和经济等等研究上有重要应用。
Deepmind 学者总结的 Reward Is Enough 观点的错误在于 reward 不是有效观察量，各种短期 reward 和长期
reward 可能指向不同、甚至相反方向，针对不同场景情况更加复杂。
追寻对称性、发现对称性破缺包括学术批判（critique），是发展人类特有智力包括语言、数学、哲学、科学等等的重要方法。数学上出现三次重要危机和
科学上形成物理学科、生命科学、智力科学三种不同范畴方法，是人类特有智力发展出独特结构和系统的关键，也是理解现在智力研究走错方向的关键。笔者从
这个角度展开研究、选择观察量、引入智力科学的二十七个入门问题。
二、智力研究的特殊性
批判（critique）一词，起源于拉丁文 criticus 和希腊文 κριτικός，跟 critical 和 judgement 有关。学术批判也是笔者新逻
辑体系中寻找逻辑临界点以及形成更高抽象、更高境界理性的关键。

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] free will [illegible]

Ştefan Odobleja [illegible] Psychologie consonantiste [illegible]

[illegible] Leukotomy [illegible] selfish gene [illegible]

[illegible] selfish gene [illegible] selfish gene [illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] judge [illegible] deadline[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] AlphaGo [illegible] dataset[illegible]

[illegible] AlphaGo [illegible] without human knowledge [illegible]

[illegible] AlphaGo Zero [illegible] AlphaGo [illegible] AlphaGo [illegible] AlphaGo [illegible]

[illegible]

[illegible]

[illegible] individualism[illegible]

[illegible]

[illegible]

[illegible]

[illegible]“[illegible]”[illegible]“[illegible]”[illegible]

[illegible]

[illegible]

[illegible]

[illegible] e [illegible] $\exp(x)$ [illegible]: $d(\exp(x))/dx = \exp(x)$[illegible]

[illegible]“[illegible]”[illegible]“[illegible]”[illegible]

[illegible]

[illegible]

[illegible]

[illegible]“[illegible]”[illegible]“[illegible]”[illegible]

[illegible]AlphoGo [illegible]

[illegible]

[illegible]The Selfish Gene[illegible] The Immortal Gene[illegible]

[illegible] word-embedding vector space [illegible] Nash Embedding Theorems [illegible]

[illegible]

[illegible]

[illegible]

[illegible]

语言和哲学上形成概念和思想、数学和科学上提出 conjecture，甚至 Grigori Perelman 证明 Poincaré conjecture 的
论文，都不是完全的逻辑过程。科学上设计验证实验，需要更复杂的思维过程。
提出有价值的 conjecture，甚至比证明单个定理更重要。有价值的 conjecture，不但可以避免散漫的搜索，而且帮助形成重要直觉，可以引出
更多有价值的发现，是应对思维和智力上熵定律趋势的重要方法。计算机能否提出有价值的 conjecture？能否证明 AlphaGo 系列可达到围棋的
最优解？如果不能，能否找出 AlphaGo 系列的问题？
严谨排除人类特有智力后，计算机的数学能力跟人类数学家相比如何？科学方法超出了计算机的能力，在应对思维和智力的熵定律趋势上能带来什么优势？如何
在计算机实验中严谨地排除人类特有智力？
研究人类特有智力和人工智能中的问题，需要研究智力进化的关键、应对熵定律趋势的各种方法以及自由意志和人类直觉思维的作用。这可以从研究语言、数学、
哲学、科学的发展着手。
如果没有毕达哥拉斯追寻对称性，也不会有希帕索斯发现无理数，更不会有后来的分析数学和现代科学。在科学上，伽利略、波义耳、达尔文开创了更严谨的追
寻对称性和发现对称性破缺的进程。而希帕索斯和伽利略揭示真相，对学术和文明的继续发展极为关键。
这些理论、方法、和思维模式的发展，为分析深度学习和其它计算机模型的局限性，提供了极为重要的线索。需要对此作系统的研究，形成新的逻辑、理论体系、
和实验方法。
科学上的对称性破缺，可以从伽利略相对性、狭义相对性、广义相对性的发展，从波义耳元素理论、拉瓦锡元素表、门捷列夫元素周期表、基本粒子标准模型的
发展、以及超对称假设的对错等等，去理解。拉瓦锡元素表有些低级错误，也有些错误却能带来重要价值。
生命科学的基础：演化理论，本身就是生物分类法的对称性破缺。由于人类特有智力的高度不稳定性，智力科学需要新的观察量和逻辑、理论、实验方法来处理
更复杂的对称性破缺。
古人把没有共同逻辑基础的《周易》、《老子》、《庄子》“融合”成玄学。玄学体系下虽不乏才智之士，但汉文明很快崩溃。玄学讲究高深莫测和高雅，以气质
上的淡泊和不屑一顾使辩论对方插不进话或说什么话都显得多余或不合适或破坏格调，从而不用严谨的辩论和公正的实验就在精神上胜过对方。
由于深度学习无法分析观察量的有效性以及数据的逻辑相容性和有效性，受深度学习影响的语言和思维也会类似玄学化，丧失逻辑严谨性，导致诡辩盛行。
Demis Hassabis 认为他们的技术有成为 a meta-solution to any problem 的潜力，还有人认为人类正在扮演上帝的角色，
就有排斥辩论和实验的高深莫测。受深度学习影响的计算机作文和计算机辩论的语言和思维也会类似玄学化。
玄学化实际上是熵定律在智力退化上的一种体现。所以需要从物理科学、生命科学、智力科学等等各种角度寻找应对熵定律的各种方法、发展更好的逻辑体系。
对三次重要数学危机深入研究，可以抽象出智力上的一些关键特征和结构，以此剖析数学研究的三种层次。计算机实际上只能从事第一数学层次的一部分工作。
大部分数学家做第一或第二数学层次的工作，在第三数学层次遭遇瓶颈。智力科学研究可以带动第三数学层次的研究。
笔者整理了二十七个智力研究入门问题，以帮助理解智力科学在观测量和逻辑、理论、实验方法上的新要求。无理数的发现和第一次数学危机引发了人类特有智
力发展中的重要问题。计算机和数字化时代暴露了这个问题实际上没有真正解决。本文的入门问题从较基本的数的语义和结构分析开始，逐步发展出各种智力原
理以及对逻辑和证伪实验的新综合。
在科学上，对称性破缺的发现，以及探索和寻找新的对称性，往往是深入研究、进一步发展的重要线索。而研究高度不稳定的智力现象，是否会帮助理解量子理
论和经典物理间的鸿沟、帮助理解现有量子理论及实验的局限性，也非常值得探索。
六、智力研究的入门问题
笔者整理的二十七个智力研究入门问题大致上可分成 A、B、C、D 四部分。问题如下：

A. 数和语言的基本结构问题：
1. 科学是否需要无理数，或对此尚不清楚、需进一步研究？如果回答不需要无理数，请给出充分理由。
2. 请完整表述无理数的语义。
3. 请找出 Chaitin's constant 语义上的问题及其对智力发展的特殊影响和争议。
4. 请构造一个例子说明不可定义数概念不稳定。
5. 请根据问题 1 - 4 抽象出语言处理的基本问题，并说明为何这是基本问题。
B. 人类直觉和统计及关联性的不同：
6. 请构造例子说明统计方法会失效并引起严重误导；并分析关联理论（relevance theory ）所带来的价值以及在抽象、分析、判断等等方面的
不足。
7. 有些群体在历史上尚处于原始部落状态时就已发展出非常复杂的构词法规则，但在当时并没有形成更灵活、表达力更丰富而又严谨的句法。请分析语言变革
和发展的需要，以及在其背后的智力结构和演变规律。
8. Grigori Perelman 的 Poincaré conjecture 论文并非用形式逻辑写成。请问除了形式逻辑，数学研究和证明在本质上还需
要什么重要思维方法？请尽可能用数学语言完整描述这些思维方法，并说明语言、数学、哲学、科学研究中所用直觉各自需要什么严谨规范。
9. Demis Hassabis 把 AlphaGo 系列中的智能称为 intuition。请问数学研究中所用到的 intuition 方法跟 Demis
Hassabis 所说的 AlphaGo 系列中的 intuition 有什么区别，类似 AlphaGo 系列的技术有没有可能成为 a meta-solution to any problem？
10. 请问 AlphaGo 系列是否如其 Nature 论文所说已达到 superhuman performance，是否是围棋的最优解？请证明或
设计实验验证你的答案，并说明验证普适性或广适性判断需要满足什么条件。
C. 广义的语言、人性、经济和自由意志：
11. 中国古代诗词经历了格律、词牌、曲牌的形成、繁荣、然后被打破，欧洲近代音乐也经历了各种 form 的形成、繁荣、然后被打破。请问：这种范式
的形成、繁荣、然后被打破，背后有什么智力规律在驱动、起了什么作用、其适用有什么特殊限制？
12. 贝多芬第五交响乐一开始的 motif 被广泛解释成命运的叩门声。但也有说法说是黄鹂叫声。请问能不能否定黄鹂叫声的说法。如果能，请说明充分
的理由。如果不能，请构造黄鹂叫声说法的合理解释。
13. 瓦格纳在音乐中想要表达的“truth”是不是真的 truth，怎么来验证？为什么他的作品中一再出现用药的情节以及用药来解释情节发展？人性和
人格发展的真正逻辑是什么？可能会导致诸神的黄昏的真正问题又是什么？
14. 请从智力原理上分析 The Selfish Gene（或 The Immortal Gene）理论的问题在哪里？这跟瓦格纳的音乐有什么关系？
15. 艺术作品常用隐喻来描述人格特征，如白天鹅、黑天鹅。Freeman Dyson 的文章 Birds and Frogs 也用 birds 和
frogs 来描述研究者的思维特征。这种用借代的具体形象来暗示抽象特征的语义难以用计算机图像处理或语言处理来获得。请分析类似的各种语言手法（包括
音乐和视觉艺术手法）背后的智力特征及其规律。
16. 请分析奥地利经济学派（Austrian School of Economics）对社会的影响和其尚未解释清楚的问题，以及各种现代经济学派的

[illegible]

17. [illegible] selfish gene [illegible]

D. [illegible]**:**

18. [illegible]

19. [illegible]

20. [illegible]“[illegible]”[illegible]“[illegible]”[illegible]logical positivism [illegible] logical empiricism[illegible]

21. [illegible] Turing Machine [illegible]deterministic, probabilistic, etc.[illegible]

22. [illegible] Turing Test [illegible] SAE level 4 [illegible] level 5 [illegible]

23. [illegible] word-embedding vector space [illegible] encoder-decoder, attention, transformer, BERT [illegible]

24. [illegible] deep-learning [illegible] deep residual networks [illegible]generative adversarial networks, etc.[illegible]

25. [illegible] Universal Approximation Theorem [illegible] overfitting[illegible]underfitting [illegible] chaos phenomena [illegible]

26. [illegible]selfish gene [illegible]

27. [illegible]

[illegible]

[illegible]

[illegible]

exact 的规律。Demis Hassabis 认为会有 a meta-solution to any problem。更说明有关科学本质的素质教育有待
改进。
第一题就涉及科学的本质。笔者问过 Freeman Dyson 教授，他回答物理学不能没有无理数。笔者认为：时空在本质上是否连续，跟时空在测量上
是否连续，是不同的问题。认为科学不需要无理数非常极端，需要提供非常充足的依据才能作为学术观点，不能误导学生。
无理数是数学概念。不管第一题答案是什么，思考第二题可以帮助理解科学和数学的分歧。该题要求完整解释语义。有疏忽就答错了。即使是功底扎实的数学家，
都可能答错。
第六题表面看似乎也不难，但题中要求“并引起严重误导”。也就是说如果构造出来的例子缺乏实际意义，就不算严重误导。
古汉语仅存构词法规则残痕，但缺乏良好的语法作严谨学术分析。二十世纪初中国的新文化运动不仅仅是通俗易懂、表达力更丰富的白话文运动，也引入了较严
谨的语法和学术分析方法，包括逻辑和诡辩分析。第七题要求从智力结构演变角度分析语言发展规律。
面对深度学习、人工智能在诡辩和其它不稳定性等等方面问题，新科学、新时代所关注焦点不同于以往，现在的优势语言特征未必能保持优势，未来所需要的语
言特征也未必能从现有特征中简单选取。这需要语言的新发展，特别是学术和科学语言。但不同自然语言采用不同的演变方式和不同的演变进度，有利于人类总
体上的多样化来更好适应未来，毕竟最优秀的学者也远非全知全能。这些问题需要审慎研究和实验，不能草率下结论。
Demis Hassabis 和 AlphaGo 系列论文的错误就跟语言有关。验证方法：可以先用一般性方法测量 AlphaGo 系列跟顶级人类棋手和
其它计算机围棋之间有让几目的精确优势；然后可以设计进一步实验比较不同的实验结果、验证智力现象的高度不稳定，从中可以看出现有语言在寻找观察量、
思考设计智力实验、分析解释实验结果和作逻辑推导时的局限性。
第十题质疑的不是 AlphaGo 后继系列对 AlphaGo Master 胜绩的可重复性，而是 Nature 论文中对实验结果的错误解释。笔者也不
质疑目前计算机围棋对人类棋手的胜绩。这些棋手尚不知道人工智能的真正问题。所以需要设计实验空间扩大搜索范围作准确的验证、测量和评估。
科学需要不断地探索更大的可能空间、接受更高抽象程度的新挑战，才能发现问题、纠正错误、继续进步。Fight-or-flight 机制会造成认知上的偏差。
高科技创业和退出机制及宣布 AlphaGo Zero 的 retire，实际上是 fight-or-flight 机制，会妨害对新科学的探索。
从第十题作更深入的分析研究，可以发现人工智能的本质缺陷和智力科学的一些基本原理。自动驾驶 SAE level 5 是普适性判断，SAE level
4 是广适性判断。
第十一题涉及音乐。音乐兼有数学和人文的特征，是非常特殊和重要的语言。作曲家需要创作自由；表演者可以选择曲目；独唱、独奏、指挥允许并需要一些不
同风格的演绎，以更好表演不同曲目；交响乐队需要听从指挥；还有伴奏、重奏、协奏等等各种配合和竞演方式，有时还需要炫技。这些专职需要不同管理方式。
非音乐专业的管理更加复杂。所以，交响乐队指挥式的管理不但不适合所有的社会管理，甚至也不适合所有的音乐界管理。
第十一题和十二题也为语言处理问题提供了更广义的注释和补充。
瓦格纳的作品和现代社会的很多问题关系密切，研究这些问题的根源和发展非常重要。第十三题中提到人性和人格发展的真正逻辑，这些明显不同于形式逻辑，
却影响到社会和文明的兴衰，是理解人类特有智力的重要关键，也涉及基于计算机和人工智能的教育系统的重要缺陷和问题，关系到人类的未来。
第十四题旨在说明智力跟生命本质有关系，作为出发点可以更深入地研究人和社会。这里要求找出除了隐含牛顿世界观的误导外，The Selfish
Gene 理论更具体、深入的问题。
第十五题中的黑、白对比鲜明而戏剧化。而真实人物的人格特征更加复杂，也更难研究。笔者曾告诉 Freeman，达尔文也是 great bird，孟德尔
才是典型的 frog。牛顿其实也兼具 bird 和 frog 两方面的特征。牛顿力学跟后来的拉格朗日力学和哈密顿力学相比，有 frog 式的局限性。而牛顿
的炼金术研究，明显是 frog 式的失败。伽利略和爱因斯坦则是典型的 bird。
Freeman 自认为是 frog，而这篇 Birds and Frogs 试图体现 bird 式的全局观。当然，Freeman 文中观点未必都准确。但
他至少对 bird 的视野还感兴趣。做 frog 式工作的人多，所以 bird 容易受排挤。如果现在连 Freeman 这样的人都很难找到，科学研究会出现

严重问题也不奇怪。
第十六题的经济学问题跟人格特征、思维模式关系密切，影响人类的生存和发展。在人类特有智力理解上的缺失和研究方法上的不足，会造成经济学研究的严重
偏差。不同行业需要不同文化。风险投资允许十分之七失败和十分之一成功那种计算，在经济学上有什么局限性，不适当地应用在教育、医疗、科学研究等等方
面会有什么问题，需要认真研究。
笔者不主张所有人都信奉奥地利经济学派。但有些影响经济发展的重要因素需要认真研究，不同经济制度和理念的优缺点及互补价值需要仔细分析，需要研究新
的经济和科技。
第十七题涉及社会和文明的组织和基础。人类教育、研究和其他社会管理，包括奖惩等等手段、方法，实际上均基于自由意志存在的假设。管理者有自由意志，
才可以选择不同管理方法、达成不同管理目标。本题仅讨论自由意志是否存在及可能的特征、规律，并不主张唯自由意志观点。
现有脑和神经系统实验仅能表明这些实验不足以寻找和分辨自由意志的特征，根本不能证明自由意志不存在。这再次说明准确使用语言、正确解释实验结果非常
重要。
第十八题回到科学研究的基础：语言问题。恰当地定义人类特有智力可以帮助设计实验来比较人类特有智力和人工智能技术，为智力研究建立科学基础。科学发
展史上，抽象了静止和运动所形成的惯性概念为经典物理和科学的建立起了关键作用。智力科学也需要更严谨的新概念来形成和发展智力原理。当然这些概念的
严谨性也需要不断发展。
第十九题中提到的诡辩，关系到司法、教育、研究、医疗等等重要问题。如果诡辩的危害无法抑制，足以瓦解科学和文明的基础。
第二十题涉及智力科学研究的几个要点。不可定义数显示语言甚至无法描述绝大部分数。哥德尔不完备定理则从结构上、系统上证明逻辑和数学语言的局限性，
包括编程语言。而“道可道，非常道；名可名，非常名”指出所有语言都有局限性；因此所有基于语言的学术也有局限性；但这只是来自直觉的哲学判断，既没有
给出证明，也没有提出解决方法。
把“道”解释成 natural law 明显错误。哥德尔不完备定理并非讨论 natural law。非欧几何已显示数学并非 natural law。而自由意
志是否存在，如果存在，跟“道”又有什么关系？
有人认为所有对称性都会破缺。哥德尔不完备定理当然正确，但这是否常道？“道可道，非常道；名可名，非常名”本身又是不是常道？如果不是，即有常道可道，
则有什么不变原理可以作为科学的严谨基石？这对倍受观察量选择不当、实验可重复性等等问题困扰的生命和智力科学研究而言非常重要。需要找出充分的逻辑
和原理依据。科学需要语言的严谨发展，需要不断继续的开放式进程，随实验条件的改进不断接受新挑战。
科学基于证伪。需要找出基于证伪实验的证实方法，综合证伪和证实。这是发展智力科学和文明继续进步的关键，是笔者的主要研究，跟第二十七问题有关。也
欢迎质疑和挑战。
第二十一题讨论计算机和人工智能的数学基础问题。
基于更广泛的问题和研究，第二十二、二十三题回到人工智能的测试理论和方法。需要针对现有人工智能以及测试方法在分析和验证智力科学原理上的不足，包
括深度学习和数据集合测试方法的不足，来做具体分析、研究和改进，把智力研究提升到科学的严谨标准。
语言是思维的工具。第二十三题要求找出计算机（包括量子计算机）在思维的表达、发展、验证上的缺陷。改进测试方法，并不是要吹毛求疵、找极端的问题。
而是要寻找重要问题，包括那些在日常生活中出现频率虽不高、却会影响科学和文明兴衰的普适性问题。
第二十三题至第二十五题也要求广泛讨论各种深度学习和人工神经网络的结构问题。对这些问题的深入研究，可以寻找各种计算机和人工智能结构的局限性的较
精确上限。
第二十六题要求对科学和文明兴衰背后的机制作更深入的分析研究。这个问题可以从科学语言和数学语言的不同着手，从物理科学、生命科学、智力科学各自使
用及未来需要的语言特征和思想方法深入研究，来寻找线索。

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

1[illegible] AlphaGo [illegible] AlphaGo Zero [illegible] AlphaGo [illegible]

2[illegible]

3[illegible]

4[illegible]

[illegible] 1) [illegible] 2) [illegible] 3) [illegible] 4) [illegible] 1) [illegible] 2) [illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] common core [illegible] global picture[illegible]

[illegible] common core [illegible]

[illegible]